

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/088, 567 A  
Source: IFWO  
Date Processed by STIC: 03/14/2006

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/10/088,567A

TIME: 10:54:35

Input Set : A:\31671178.APP

Output Set: N:\CRF4\03142006\J088567A.raw

```

3 <110> APPLICANT: AKIRA, SHIZUO
4   HEMMI, HIROAKI
6 <120> TITLE OF INVENTION: RECEPTOR PROTEINS SPECIFICALLY RECOGNIZING BACTERIAL
7   DNA
9 <130> FILE REFERENCE: 31671-178057
11 <140> CURRENT APPLICATION NUMBER: 10/088,567A
12 <141> CURRENT FILING DATE: 2002-03-19
14 <150> PRIOR APPLICATION NUMBER: PCT/JP01/04731
15 <151> PRIOR FILING DATE: 2001-06-05
17 <150> PRIOR APPLICATION NUMBER: JP 2000-219652
18 <151> PRIOR FILING DATE: 2000-07-19
20 <160> NUMBER OF SEQ ID NOS: 9
22 <170> SOFTWARE: PatentIn Ver. 3.3
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 3257
26 <212> TYPE: DNA
27 <213> ORGANISM: Homo sapiens
29 <220> FEATURE:
30 <221> NAME/KEY: CDS
31 <222> LOCATION: (107)..(3205)
33 <400> SEQUENCE: 1
34 ccgctgctgc ccctgtggga agggacctcg agtgtgaagc atccttccct gtagctgctg 60
36 tccagtctgc ccgccagacc ctctggagaa gcccctgccc cccagc atg ggt ttc 115
37                                     Met Gly Phe
38                                     1
40 tgc cgc agc gcc ctg cac ccg ctg tct ctc ctg gtg cag gcc atc atg 163
41 Cys Arg Ser Ala Leu His Pro Leu Ser Leu Leu Val Gln Ala Ile Met
42   5                               10                               15
44 ctg gcc atg acc ctg gcc ctg ggt acc ttg cct gcc ttc cta ccc tgt 211
45 Leu Ala Met Thr Leu Ala Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys
46 20                               25                               30                               35
48 gag ctc cag ccc cac ggc ctg gtg aac tgc aac tgg ctg ttc ctg aag 259
49 Glu Leu Gln Pro His Gly Leu Val Asn Cys Asn Trp Leu Phe Leu Lys
50                               40                               45                               50
52 tct gtg ccc cac ttc tcc atg gca gca ccc cgt ggc aat gtc acc agc 307
53 Ser Val Pro His Phe Ser Met Ala Ala Pro Arg Gly Asn Val Thr Ser
54                               55                               60                               65
56 ctt tcc ttg tcc tcc aac cgc atc cac cac ctc cat gat tct gac ttt 355
57 Leu Ser Leu Ser Ser Asn Arg Ile His His Leu His Asp Ser Asp Phe
58   70                               75                               80
60 gcc cac ctg ccc agc ctg cgg cat ctc aac ctc aag tgg aac tgc ccg 403
61 Ala His Leu Pro Ser Leu Arg His Leu Asn Leu Lys Trp Asn Cys Pro
62   85                               90                               95

```

## RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/10/088,567A

TIME: 10:54:35

Input Set : A:\31671178.APP

Output Set: N:\CRF4\03142006\J088567A.raw

```

64 ccg gtt ggc ctc agc ccc atg cac ttc ccc tgc cac atg acc atc gag 451
65 Pro Val Gly Leu Ser Pro Met His Phe Pro Cys His Met Thr Ile Glu
66 100 105 110 115
68 ccc agc acc ttc ttg gct gtg ccc acc ctg gaa gag cta aac ctg agc 499
69 Pro Ser Thr Phe Leu Ala Val Pro Thr Leu Glu Glu Leu Asn Leu Ser
70 120 125 130
72 tac aac aac atc atg act gtg cct gcg ctg ccc aaa tcc ctc ata tcc 547
73 Tyr Asn Asn Ile Met Thr Val Pro Ala Leu Pro Lys Ser Leu Ile Ser
74 135 140 145
76 ctg tcc ctc agc cat acc aac atc ctg atg cta gac tct gcc agc ctc 595
77 Leu Ser Leu Ser His Thr Asn Ile Leu Met Leu Asp Ser Ala Ser Leu
78 150 155 160
80 gcc ggc ctg cat gcc ctg cgc ttc cta ttc atg gac ggc aac tgt tat 643
81 Ala Gly Leu His Ala Leu Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr
82 165 170 175
84 tac aag aac ccc tgc agg cag gca ctg gag gtg gcc ccg ggt gcc ctc 691
85 Tyr Lys Asn Pro Cys Arg Gln Ala Leu Glu Val Ala Pro Gly Ala Leu
86 180 185 190 195
88 ctt ggc ctg ggc aac ctc acc cac ctg tca ctc aag tac aac aac ctc 739
89 Leu Gly Leu Gly Asn Leu Thr His Leu Ser Leu Lys Tyr Asn Asn Leu
90 200 205 210
92 act gtg gtg ccc cgc aac ctg cct tcc agc ctg gag tat ctg ctg ttg 787
93 Thr Val Val Pro Arg Asn Leu Pro Ser Ser Leu Glu Tyr Leu Leu Leu
94 215 220 225
96 tcc tac aac cgc atc gtc aaa ctg gcg cct gag gac ctg gcc aat ctg 835
97 Ser Tyr Asn Arg Ile Val Lys Leu Ala Pro Glu Asp Leu Ala Asn Leu
98 230 235 240
100 acc gcc ctg cgt gtg ctc gat gtg ggc gga aat tgc cgc cgc tgc gac 883
101 Thr Ala Leu Arg Val Leu Asp Val Gly Gly Asn Cys Arg Arg Cys Asp
102 245 250 255
104 cac gct ccc aac ccc tgc atg gag tgc cct cgt cac ttc ccc cag cta 931
105 His Ala Pro Asn Pro Cys Met Glu Cys Pro Arg His Phe Pro Gln Leu
106 260 265 270 275
108 cat ccc gat acc ttc agc cac ctg agc cgt ctt gaa ggc ctg gtg ttg 979
109 His Pro Asp Thr Phe Ser His Leu Ser Arg Leu Glu Gly Leu Val Leu
110 280 285 290
112 aag gac agt tct ctc tcc tgg ctg aat gcc agt tgg ttc cgt ggg ctg 1027
113 Lys Asp Ser Ser Leu Ser Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu
114 295 300 305
116 gga aac ctc cga gtg ctg gac ctg agt gag aac ttc ctc tac aaa tgc 1075
117 Gly Asn Leu Arg Val Leu Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys
118 310 315 320
120 atc act aaa acc aag gcc ttc cag ggc cta aca cag ctg cgc aag ctt 1123
121 Ile Thr Lys Thr Lys Ala Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu
122 325 330 335
124 aac ctg tcc ttc aat tac caa aag agg gtg tcc ttt gcc cac ctg tct 1171
125 Asn Leu Ser Phe Asn Tyr Gln Lys Arg Val Ser Phe Ala His Leu Ser
126 340 345 350 355
128 ctg gcc cct tcc ttc ggg agc ctg gtc gcc ctg aag gag ctg gac atg 1219

```

## RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/10/088,567A

TIME: 10:54:35

Input Set : A:\31671178.APP

Output Set: N:\CRF4\03142006\J088567A.raw

129	Leu	Ala	Pro	Ser	Phe	Gly	Ser	Leu	Val	Ala	Leu	Lys	Glu	Leu	Asp	Met	
130					360					365					370		
132	cac	ggc	atc	ttc	ttc	cgc	tca	ctc	gat	gag	acc	acg	ctc	cgg	cca	ctg	1267
133	His	Gly	Ile	Phe	Phe	Arg	Ser	Leu	Asp	Glu	Thr	Thr	Leu	Arg	Pro	Leu	
134				375					380					385			
136	gcc	cgc	ctg	ccc	atg	ctc	cag	act	ctg	cgt	ctg	cag	atg	aac	ttc	atc	1315
137	Ala	Arg	Leu	Pro	Met	Leu	Gln	Thr	Leu	Arg	Leu	Gln	Met	Asn	Phe	Ile	
138				390					395					400			
140	aac	cag	gcc	cag	ctc	ggc	atc	ttc	agg	gcc	ttc	cct	ggc	ctg	cgc	tac	1363
141	Asn	Gln	Ala	Gln	Leu	Gly	Ile	Phe	Arg	Ala	Phe	Pro	Gly	Leu	Arg	Tyr	
142				405					410					415			
144	gtg	gac	ctg	tcg	gac	aac	cgc	atc	agc	gga	gct	tcg	gag	ctg	aca	gcc	1411
145	Val	Asp	Leu	Ser	Asp	Asn	Arg	Ile	Ser	Gly	Ala	Ser	Glu	Leu	Thr	Ala	
146	420					425								430		435	
148	acc	atg	ggg	gag	gca	gat	gga	ggg	gag	aag	gtc	tgg	ctg	cag	cct	ggg	1459
149	Thr	Met	Gly	Glu	Ala	Asp	Gly	Gly	Glu	Lys	Val	Trp	Leu	Gln	Pro	Gly	
150					440					445						450	
152	gac	ctt	gct	ccg	gcc	cca	gtg	gac	act	ccc	agc	tct	gaa	gac	ttc	agg	1507
153	Asp	Leu	Ala	Pro	Ala	Pro	Val	Asp	Thr	Pro	Ser	Ser	Glu	Asp	Phe	Arg	
154					455					460						465	
156	ccc	aac	tgc	agc	acc	ctc	aac	ttc	acc	ttg	gat	ctg	tca	cgg	aac	aac	1555
157	Pro	Asn	Cys	Ser	Thr	Leu	Asn	Phe	Thr	Leu	Asp	Leu	Ser	Arg	Asn	Asn	
158					470					475						480	
160	ctg	gtg	acc	gtg	cag	ccg	gag	atg	ttt	gcc	cag	ctc	tcg	cac	ctg	cag	1603
161	Leu	Val	Thr	Val	Gln	Pro	Glu	Met	Phe	Ala	Gln	Leu	Ser	His	Leu	Gln	
162										485						490	
164	tgc	ctg	cgc	ctg	agc	cac	aac	tgc	atc	tcg	cag	gca	gtc	aat	ggc	tcc	1651
165	Cys	Leu	Arg	Leu	Ser	His	Asn	Cys	Ile	Ser	Gln	Ala	Val	Asn	Gly	Ser	
166	500						505									510	
168	cag	ttc	ctg	ccg	ctg	acc	ggt	ctg	cag	gtg	cta	gac	ctg	tcc	cac	aat	1699
169	Gln	Phe	Leu	Pro	Leu	Thr	Gly	Leu	Gln	Val	Leu	Asp	Leu	Ser	His	Asn	
170						520										525	
172	aag	ctg	gac	ctc	tac	cac	gag	cac	tca	ttc	acg	gag	cta	cca	cga	ctg	1747
173	Lys	Leu	Asp	Leu	Tyr	His	Glu	His	Ser	Phe	Thr	Glu	Leu	Pro	Arg	Leu	
174						535										540	
176	gag	gcc	ctg	gac	ctc	agc	tac	aac	agc	cag	ccc	ttt	ggc	atg	cag	ggc	1795
177	Glu	Ala	Leu	Asp	Leu	Ser	Tyr	Asn	Ser	Gln	Pro	Phe	Gly	Met	Gln	Gly	
178										550						555	
180	gtg	ggc	cac	aac	ttc	agc	ttc	gtg	gct	cac	ctg	cgc	acc	ctg	cgc	cac	1843
181	Val	Gly	His	Asn	Phe	Ser	Phe	Val	Ala	His	Leu	Arg	Thr	Leu	Arg	His	
182										565						570	
184	ctc	agc	ctg	gcc	cac	aac	aac	atc	cac	agc	caa	gtg	tcc	cag	cag	ctc	1891
185	Leu	Ser	Leu	Ala	His	Asn	Asn	Ile	His	Ser	Gln	Val	Ser	Gln	Gln	Leu	
186	580															585	
188	tgc	agt	acg	tcg	ctg	cgg	gcc	ctg	gac	ttc	agc	ggc	aat	gca	ctg	ggc	1939
189	Cys	Ser	Thr	Ser	Leu	Arg	Ala	Leu	Asp	Phe	Ser	Gly	Asn	Ala	Leu	Gly	
190										600						605	
192	cat	atg	tgg	gcc	gag	gga	gac	ctc	tat	ctg	cac	ttc	ttc	caa	ggc	ctg	1987
193	His	Met	Trp	Ala	Glu	Gly	Asp	Leu	Tyr	Leu	His	Phe	Phe	Gln	Gly	Leu	

## RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/10/088,567A

TIME: 10:54:35

Input Set : A:\31671178.APP

Output Set: N:\CRF4\03142006\J088567A.raw

194	615	620	625	
196	agc ggt ttg atc tgg ctg gac ttg tcc cag aac cgc ctg cac acc ctc	2035		
197	Ser Gly Leu Ile Trp Leu Asp Leu Ser Gln Asn Arg Leu His Thr Leu			
198	630	635	640	
200	ctg ccc caa acc ctg cgc aac ctc ccc aag agc cta cag gtg ctg cgt	2083		
201	Leu Pro Gln Thr Leu Arg Asn Leu Pro Lys Ser Leu Gln Val Leu Arg			
202	645	650	655	
204	ctc cgt gac aat tac ctg gcc ttc ttt aag tgg tgg agc ctc cac ttc	2131		
205	Leu Arg Asp Asn Tyr Leu Ala Phe Phe Lys Trp Trp Ser Leu His Phe			
206	660	665	670	675
208	ctg ccc aaa ctg gaa gtc ctc gac ctg gca gga aac cag ctg aag gcc	2179		
209	Leu Pro Lys Leu Glu Val Leu Asp Leu Ala Gly Asn Gln Leu Lys Ala			
210	680	685	690	
212	ctg acc aat ggc agc ctg cct gct ggc acc cgg ctc cgg agg ctg gat	2227		
213	Leu Thr Asn Gly Ser Leu Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp			
214	695	700	705	
216	gtc agc tgc aac agc atc agc ttc gtg gcc ccc ggc ttc ttt tcc aag	2275		
217	Val Ser Cys Asn Ser Ile Ser Phe Val Ala Pro Gly Phe Phe Ser Lys			
218	710	715	720	
220	gcc aag gag ctg cga gag ctc aac ctt agc gcc aac gcc ctc aag aca	2323		
221	Ala Lys Glu Leu Arg Glu Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr			
222	725	730	735	
224	gtg gac cac tcc tgg ttt ggg ccc ctg gcg agt gcc ctg caa ata cta	2371		
225	Val Asp His Ser Trp Phe Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu			
226	740	745	750	755
228	gat gta agc gcc aac cct ctg cac tgc gcc tgt ggg gcg gcc ttt atg	2419		
229	Asp Val Ser Ala Asn Pro Leu His Cys Ala Cys Gly Ala Ala Phe Met			
230	760	765	770	
232	gac ttc ctg ctg gag gtg cag gct gcc gtg ccc ggt ctg ccc agc cgg	2467		
233	Asp Phe Leu Leu Glu Val Gln Ala Ala Val Pro Gly Leu Pro Ser Arg			
234	775	780	785	
236	gtg aag tgt ggc agt ccg ggc cag ctc cag ggc ctc agc atc ttt gca	2515		
237	Val Lys Cys Gly Ser Pro Gly Gln Leu Gln Gly Leu Ser Ile Phe Ala			
238	790	795	800	
240	cag gac ctg cgc ctc tgc ctg gat gag gcc ctc tcc tgg gac tgt ttc	2563		
241	Gln Asp Leu Arg Leu Cys Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe			
242	805	810	815	
244	gcc ctc tgc ctg ctg gct gtg gct ctg ggc ctg ggt gtg ccc atg ctg	2611		
245	Ala Leu Ser Leu Leu Ala Val Ala Leu Gly Leu Gly Val Pro Met Leu			
246	820	825	830	835
248	cat cac ctc tgt ggc tgg gac ctc tgg tac tgc ttc cac ctg tgc ctg	2659		
249	His His Leu Cys Gly Trp Asp Leu Trp Tyr Cys Phe His Leu Cys Leu			
250	840	845	850	
252	gcc tgg ctt ccc tgg cgg ggg cgg caa agt ggg cga gat gag gat gcc	2707		
253	Ala Trp Leu Pro Trp Arg Gly Arg Gln Ser Gly Arg Asp Glu Asp Ala			
254	855	860	865	
256	ctg ccc tac gat gcc ttc gtg gtc ttc gac aaa acg cag agc gca gtg	2755		
257	Leu Pro Tyr Asp Ala Phe Val Val Phe Asp Lys Thr Gln Ser Ala Val			
258	870	875	880	

## RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/10/088,567A

TIME: 10:54:35

Input Set : A:\31671178.APP

Output Set: N:\CRF4\03142006\J088567A.raw

```

260 gca gac tgg gtg tac aac gag ctt cgg ggg cag ctg gag gag tgc cgt 2803
261 Ala Asp Trp Val Tyr Asn Glu Leu Arg Gly Gln Leu Glu Glu Cys Arg
262      885                      890                      895
264 ggg cgc tgg gca ctc cgc ctg tgc ctg gag gaa cgc gac tgg ctg cct 2851
265 Gly Arg Trp Ala Leu Arg Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro
266 900                      905                      910                      915
268 ggc aaa acc ctc ttt gag aac ctg tgg gcc tcg gtc tat ggc agc cgc 2899
269 Gly Lys Thr Leu Phe Glu Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg
270                      920                      925                      930
272 aag acg ctg ttt gtg ctg gcc cac acg gac cgg gtc agt ggt ctc ttg 2947
273 Lys Thr Leu Phe Val Leu Ala His Thr Asp Arg Val Ser Gly Leu Leu
274                      935                      940                      945
276 cgc gcc agc ttc ctg ctg gcc cag cag cgc ctg ctg gag gac cgc aag 2995
277 Arg Ala Ser Phe Leu Leu Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys
278      950                      955                      960
280 gac gtc gtg gtg ctg gtg atc ctg agc cct gac ggc cgc cgc tcc cgc 3043
281 Asp Val Val Val Leu Val Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg
282      965                      970                      975
284 tac gtg cgg ctg cgc cag cgc ctc tgc cgc cag agt gtc ctc ctc tgg 3091
285 Tyr Val Arg Leu Arg Gln Arg Leu Cys Arg Gln Ser Val Leu Leu Trp
286 980                      985                      990                      995
288 ccc cac cag ccc agt ggt cag cgc agc ttc tgg gcc cag ctg ggc atg 3139
289 Pro His Gln Pro Ser Gly Gln Arg Ser Phe Trp Ala Gln Leu Gly Met
290                      1000                      1005                      1010
292 gcc ctg acc agg gac aac cac cac ttc tat aac cgg aac ttc tgc cag 3187
293 Ala Leu Thr Arg Asp Asn His His Phe Tyr Asn Arg Asn Phe Cys Gln
294                      1015                      1020                      1025
296 gga ccc acg gcc gaa tag ccgtgagccg gaatcctgca cgggtgccacc 3235
297 Gly Pro Thr Ala Glu
298      1030
300 tccacactca cctcacctct gc 3257
303 <210> SEQ ID NO: 2
304 <211> LENGTH: 1032
305 <212> TYPE: PRT
306 <213> ORGANISM: Homo sapiens
308 <400> SEQUENCE: 2
309 Met Gly Phe Cys Arg Ser Ala Leu His Pro Leu Ser Leu Leu Val Gln
310 1 5 10 15
312 Ala Ile Met Leu Ala Met Thr Leu Ala Leu Gly Thr Leu Pro Ala Phe
313 20 25 30
315 Leu Pro Cys Glu Leu Gln Pro His Gly Leu Val Asn Cys Asn Trp Leu
316 35 40 45
318 Phe Leu Lys Ser Val Pro His Phe Ser Met Ala Ala Pro Arg Gly Asn
319 50 55 60
321 Val Thr Ser Leu Ser Leu Ser Ser Asn Arg Ile His His Leu His Asp
322 65 70 75 80
324 Ser Asp Phe Ala His Leu Pro Ser Leu Arg His Leu Asn Leu Lys Trp
325 85 90 95
327 Asn Cys Pro Pro Val Gly Leu Ser Pro Met His Phe Pro Cys His Met

```

**VERIFICATION SUMMARY**

DATE: 03/14/2006

PATENT APPLICATION: US/10/088,567A

TIME: 10:54:36

Input Set : A:\31671178.APP

Output Set: N:\CRF4\03142006\J088567A.raw